



## Statistical evaluation of characteristic earthquakes in the frequency-magnitude distributions of Sumatra and other subduction zone regions

Naylor, M., Greenhough, J., McCloskey, J., A.F. Bell, A., & Main, I. (2009). Statistical evaluation of characteristic earthquakes in the frequency-magnitude distributions of Sumatra and other subduction zone regions. *Geophysical Research Letters*, 36(L20303). <https://doi.org/10.1029/2009GL040460>

[Link to publication record in Ulster University Research Portal](#)

**Published in:**  
Geophysical Research Letters

**Publication Status:**  
Published (in print/issue): 17/10/2009

**DOI:**  
[10.1029/2009GL040460](https://doi.org/10.1029/2009GL040460)

**Document Version**  
Publisher's PDF, also known as Version of record

**General rights**  
Copyright for the publications made accessible via Ulster University's Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**  
The Research Portal is Ulster University's institutional repository that provides access to Ulster's research outputs. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact [pure-support@ulster.ac.uk](mailto:pure-support@ulster.ac.uk).

# Statistical evaluation of characteristic earthquakes in the frequency-magnitude distributions of Sumatra and other subduction zone regions

M. Naylor,<sup>1</sup> J. Greenhough,<sup>1</sup> J. McCloskey,<sup>2</sup> A.F. Bell,<sup>1</sup> and I.G. Main<sup>1</sup>

Received 7 August 2009; revised 31 August 2009; accepted 2 September 2009; published 17 October 2009.

[1] If subduction zone earthquakes conform to a characteristic model, in which persistent segments fail at predictable stress levels due to the steady accumulation of tectonic loading, historical seismicity may constrain the occurrence of future events. We test this model for earthquakes on the Sumatra-Andaman megathrust and other subduction zones using frequency-magnitude distributions. Using simulations, we show that Poisson confidence intervals correctly account for the counting errors of histogram data. These confidence intervals demonstrate that we cannot reject the Gutenberg-Richter distribution in favor of a characteristic model in any of the real catalogues tested. A visual bias in power-law count data at high magnitudes, combined with a sample bias for large earthquakes, is sufficient to explain candidate characteristic events. This result implies that historical earthquakes are likely poor models for future events and that Monte Carlo simulations will provide a better assessment of earthquake and associated hazards. **Citation:** Naylor, M., J. Greenhough, J. McCloskey, A. F. Bell, and I. G. Main (2009), Statistical evaluation of characteristic earthquakes in the frequency-magnitude distributions of Sumatra and other subduction zone regions, *Geophys. Res. Lett.*, 36, L20303, doi:10.1029/2009GL040460.

## 1. Introduction

[2] There exists an established systematic relation between an earthquake's magnitude,  $M$  and its frequency of occurrence in any given region and time period,  $N$  [e.g., Gutenberg and Richter, 1954]. The Gutenberg–Richter (GR) relation can be expressed as,

$$\log_{10} N = a - bM \quad (1)$$

where  $a$  and  $b$  are constants, with  $b$  typically close to 1. Given the logarithmic relation between an earthquake's magnitude and its energy or moment, the GR relation reflects a power-law relation between frequency and size.

[3] If the GR relation constituted an exact description of the frequency of earthquake magnitudes, the underlying power-law implies that a larger event will always be observed if we wait sufficiently long. However, the Earth's finite size enforces an upper limit on the size of earthquakes and hence a roll-off to the GR distribution at high magnitudes, for example, an exponential tail (a Gamma distribu-

tion) [Main and Burton, 1984]; the global catalogue does not yet demonstrate this [Main *et al.*, 2008].

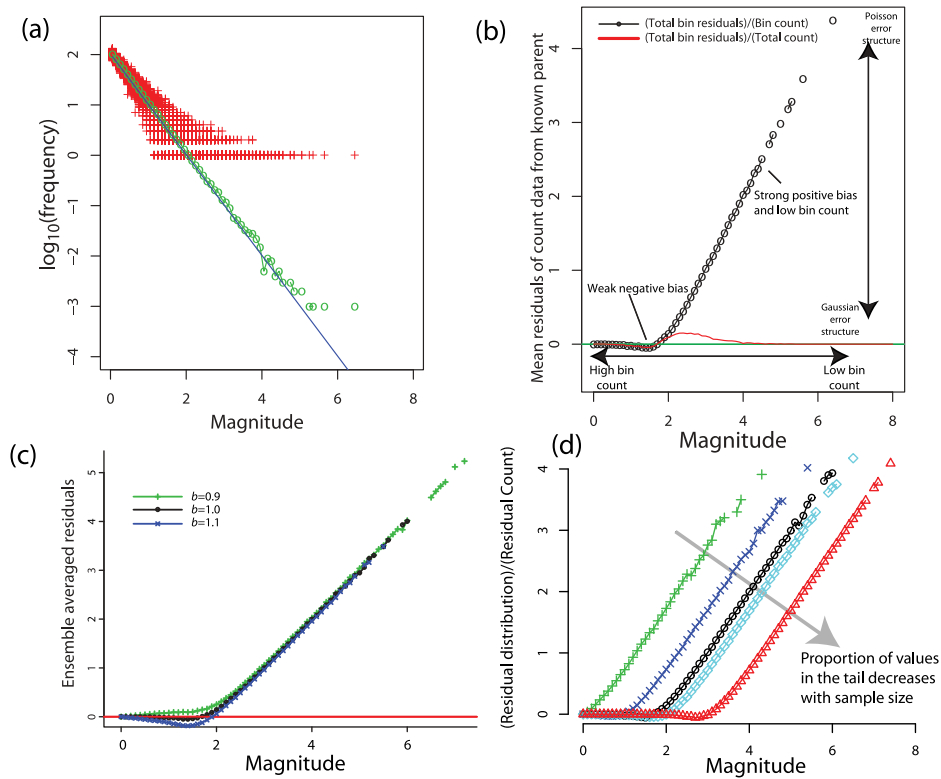
[4] In addition to evidence for GR and Gamma distributions, geological and palaeo-seismological data has been used to argue that “characteristic” earthquakes may result from repeated rupture of the same patch of fault [Wesnowsky, 1994]. Characteristic-type behaviour also indicates a degree of predictability in the system, particularly with respect to earthquake location and magnitude. Sieh *et al.* [2008] highlight the similarities between rupture locations of palaeo-earthquakes on the Sunda megathrust and events following the 2004 Andaman Islands earthquake, and argue that this is indicative of characteristic behaviour. The characteristic earthquake model imposes strong constraints on the nature of impending events and associated hazards. The expectation of a repeat of the 1797 Menatawai Islands earthquake predicts a tsunami for the city of Padang on western Sumatra which could reach 5–6m locally [Borrero *et al.*, 2006]. A non-characteristic model for the location and size of future earthquakes forecasts a much wider range of possible tsunamis [e.g., McCloskey *et al.*, 2008]. At present the characteristic model informs evacuation planning in western Sumatra; it is clearly desirable to use independent methods to test this hypothesis. Recent large earthquakes on the Sunda megathrust do not discriminate well between these models. Whereas the 1833 earthquake would appear to have been a good model for the spatial extent and moment of the 2005 Nias earthquake [Briggs *et al.*, 2006], the 2007 events could not have been expected [Konca *et al.*, 2008].

[5] A key manifestation of the characteristic earthquake hypothesis is a greater frequency of large events than predicted by extrapolation of the GR relation, adding a ‘bump’ to the tail of the log-linear distribution. Thus frequency-magnitude histograms potentially provide a test for characteristic behaviour. However, in many cases the evidence remains qualitative [e.g., Schwartz and Coppersmith, 1984]. Where formal statistical tests are done, it is common to neglect potential biases from the effects of finite sample size and to assume that residuals are Gaussian distributed [e.g., Speidel and Mattson, 1997].

[6] We consider random samples drawn from power-law distributions to discern the true properties of residuals in power-law count data. By averaging many synthetic earthquake catalogue realizations, (1) discrete numbers of observations converge towards a continuous probability density function that approximates the parent (underlying) distribution, and (2) significant biases associated with discrete samples (real earthquake magnitude observations) are quantified. Importantly, synthetic samples drawn from power-laws are free from biases that may exist in real data, for example, magnitude saturation in seismometers.

<sup>1</sup>School of GeoSciences, University of Edinburgh, Edinburgh, UK.

<sup>2</sup>School of Environmental Sciences, University of Ulster, Ulster, UK.



**Figure 1.** Synthetic power-law frequency-magnitude data and their statistical properties are derived from 1000 randomly drawn samples each containing 500 events from a power-law distribution with exponent  $b = 1$  unless otherwise stated. (a) One thousand histograms (red crosses) are plotted, with the parent distribution offset to the average count for the lowest bin (blue line) and mean bin occupancy which averages the count data in each bin (green circles), (b) total residuals (deviation in frequency from the parent distribution in log-space) in each bin normalised by the number of events in the bin to find the mean deviation from the parent, (c) variation in residuals with exponent  $b$ , and (d) effect of sample length on residuals in 10, 100, 500, 1000, and 10000 events, left to right.

[7] The techniques described are generic to testing the consistency of any data against the null hypothesis of a power-law frequency-size distribution, with wide applicability elsewhere in geophysics, including natural hazard assessment.

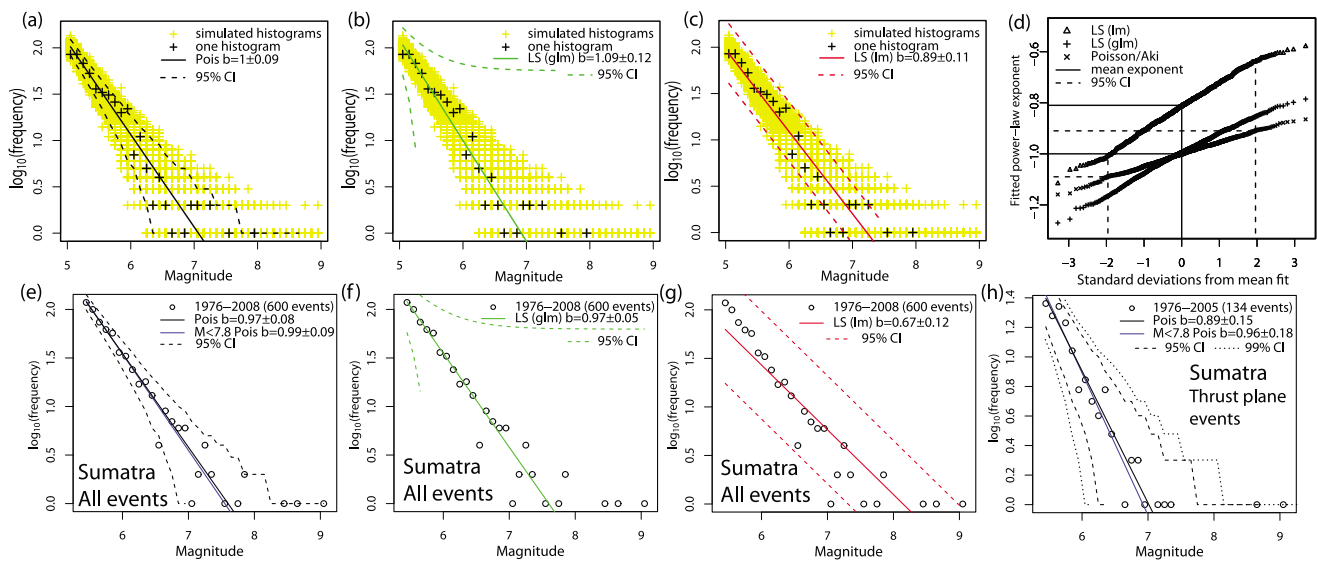
## 2. What to Expect When Sampling From Power-Laws

[8] Intuition can be misleading when assessing whether samples are, or are not, drawn from power-law distributions. Generally, the mean of an unbounded power-law distribution is not finite; it increases systematically with sample size. In contrast, a distribution that exhibits Gaussian convergence, for example, finding the mean height of the population, converges to a stable mean rapidly with increasing sample size. Counting uncertainties may also markedly affect the appearance of a log-log histogram. To demonstrate this we have generated 1000 synthetic catalogues containing 500 events drawn from a power-law energy distribution with exponent  $b = 1$ ; after taking the  $\log_{10}$  of the values drawn from the power-law distribution we generate a histogram for each synthetic catalogue with a bin width of 0.1 in log space and superimpose the histograms in Figure 1a using red crosses to highlight the spread in the samples (throughout the paper, the  $\log_{10}$  transformation is used to produce the magnitude scale from the power-law distributed energies).

The magnitude axis is proportional to the  $\log_{10}$  of the energy. The blue line is the power-law distribution from which samples are drawn with the slope equal to the exponent. Figure 1a demonstrates that sampling from a power-law distribution produces a visual bias that increases with magnitude; extreme events appear to have a frequency larger than that predicted by the parent distribution.

[9] The origin of this observational bias is clear. As histograms show count data, all of the non-empty bins contain 1 or more events ( $\log_{10} N \geq 0$ ). Some of the counts in a random sample correspond to bins where the mean frequency is less than 1. By averaging over all replicates, we can calculate the mean count in each bin (green circles in Figure 1a), which represents a catalogue with 500,000 events. These mean frequencies recover the parent distribution (blue line). Therefore, visual assessment can be misleading as to whether samples are drawn from power-law distributions; a single sample is insufficient to represent the full parent distribution (blue line). It is in this region of sparse data where candidate characteristic events lie.

[10] The average residuals for each bin are shown in Figure 1b. The residuals vary with magnitude and can be divided into three regions. At low magnitudes the number of events in each bin  $\gg 1$  and counting errors are small, approximately symmetric and approximately Gaussian about the parent distribution. For frequencies of  $\sim 1$  at intermediate magnitudes, residuals appear negatively biased



**Figure 2.** Comparison of different fitting schemes on (top) synthetic data (1000 synthetic samples drawn from a power-law distribution with exponent  $b = 1$  of 500 events) and (bottom) real data from Sumatra-Andaman (bottom row). (a–c) Frequency-magnitude plots for simulations with the corresponding confidence intervals calculated from a single sample for each of the techniques described in the text and (d) b-value estimates and standard deviations of all those estimates from the mean estimate for each technique. Sumatra-Andaman frequency-magnitude data are shown fitted with (e) Poisson residuals, (f) Gaussian residuals in linear space, (g) least squares with Gaussian residuals in log space, and (h) Poisson residuals using a subset of Sumatra-Andaman events that lie on the megathrust.

because these events are under-sampled. The errors are clearly non-Gaussian and are approximately Poisson.

[11] For the largest events when the mean frequency is below 1, all occupied bins lie above the mean and hence appear positively biased; this may account for apparent characteristic behaviour.

[12] Since the region of negative residuals is a consequence of the counting errors, its range is modified by changes in  $b$  (Figure 1c). Decreasing  $b$  decreases the slope and accentuates the negative residuals; increasing  $b$  changes the sign of the residuals in this region. A steeper parent slope results in increasing positive bias.

[13] The residual analysis is repeated for different sample sizes (Figure 1d). All curves show positive residuals at high magnitudes; the deviation occurs at greater magnitudes as the sample size increases. The gradient of the positive residuals region is the same for all sample sizes. The sample of 10 is dominated by positively biased counting errors. Above  $\sim 500$  events, the well-sampled region is resolved for smaller magnitudes, but there are still  $\sim 50\%$  of values in the positively biased tail. Increasing the sample size increases the proportion of counts  $\gg 1$  but maintains the same mean number of events in the tail (Figure 1d). Therefore, since single-sample events in the tail cannot be better resolved by waiting for more events, evidence for non-power-law distributions can lie only significantly beyond the counting errors.

### 3. Implications for Fitting Data

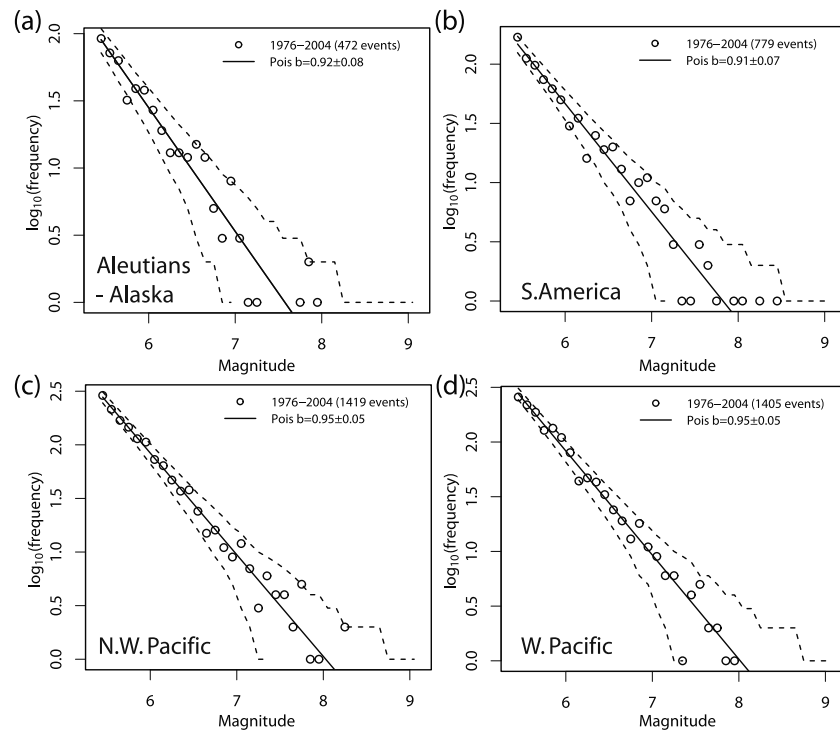
[14] Standard least squares regression, especially in the case of log counts, does not correctly model the counting errors described above [Sandri and Marzocchi, 2007]. However, there remain numerous examples where the

regression technique is either not stated or is clearly inappropriate [e.g., Leonard, 2008; Liu *et al.*, 2007; Polat *et al.*, 2008; Sayil and Osmansahim, 2008]. Using three ways to fit a power-law to the histogram, we demonstrate how different Generalised Linear Models (GLM) account for different parts of the bias. We compare (1) the least-biased (i.e., Maximum Likelihood) method using Poisson residuals (Figure 2a), with the frequently but incorrectly applied (2) Gaussian residuals (least squares in linear space) (Figure 2b), and (3) least squares with Gaussian residuals in log space (Figure 2c). It has been shown formally [Charnes *et al.*, 1976] that GLM with Poisson residuals is mathematically equivalent Aki's [1965] Maximum Likelihood technique often used for fitting earthquake histograms. We reserve a detailed discussion of each of these techniques for the auxiliary material.<sup>1</sup>

[15] Confidence intervals show regions within which, in the limit of a large number of bins, a certain fraction of bin frequencies are expected to lie; this constitutes an approximate retrospective test for a correctly specified residual distribution. To each plot (Figures 2a–2c) we have added the corresponding 95% confidence limits (dashed lines) on the 2.5 and 97.5 percentiles.

[16] Confidence intervals for the Poisson residuals (Figure 2a) correctly account for the variability in synthetic data (i.e., 95% of the counts lie within the 95% confidence intervals), unlike either of the other fits (Figures 2b–2c), and crucially capture the presence of strong positive residuals at high magnitudes. Furthermore, Figure 2d shows estimates of the power-law exponents  $b$  for each sample

<sup>1</sup>Auxiliary materials are available in the HTML. doi:10.1029/2009GL040460.



**Figure 3.** Histograms and Poisson 95% confidence limits for (a) Aleutians and Alaska: the Aleutians and Alaska; (b) South America: west coast of South America; (c) Pacific North-West: Kamchatka, the Kurils, Honshu, Hokkaido, and the Marianas; and (d) West Pacific: Kyushu, Ryuku islands, Taiwan, Luzon and the Molucca passage.

using each technique. The Poisson method is not systematically biased. The Gaussian estimate (least squares in linear space) is not biased either but suffers from greater variability in individual estimates. The least-squares in log space technique displays both systematic bias and large variability; the known  $b$  used to generate the samples barely lies within the 95% confidence intervals.

[17] Even authors who explicitly state that they use a Maximum Likelihood method often fail to specify the distribution of residuals assumed in the regression. GLM software normally defaults to Gaussian residuals and hence this hidden assumption could lie behind much suboptimal fitting in the literature. Here we stress that the true Maximum Likelihood fit is carried out using Poisson residuals, which usually requires only the setting of a flag in the GLM function. To facilitate its use we explicitly provide the ‘R’ code for fitting histograms in the auxiliary material.

#### 4. Analysis of Real Catalogues

[18] We now examine the histograms of earthquake magnitudes recorded in the Harvard CMT catalogue between 1977 and 2007 for the Sunda megathrust and comparable subduction zone regions (see auxiliary material for selection criteria) in order to test for characteristic earthquakes. More than 600  $M_w > 5.4$  events are recorded in the Sunda megathrust region. The Poisson fit and confidence intervals (see section 3) best characterise the histogram (Figures 2e–2g). The apparent ‘outliers’ at high magnitudes are entirely consistent with the power-law null hypothesis (see section 2). Least squares (in log space) produces a poor

fit to the bulk of the data at lower magnitudes (where the fit should be best) and the confidence intervals for the Gaussian fit (in linear space) are clearly too wide.

[19] We also analyse a subset of the catalogue for events whose hypocentre locations and focal mechanisms suggest they lie on the megathrust interface. These events might be more likely to display a characteristic-type distribution (Figure 2h). However, the confidence intervals on the counts are wider due to the smaller sample (Figure 1c). One event lies beyond the 95% confidence limit, but within the 99% confidence limit (the 99% limit for a count of 1 is at magnitude 9.5). Since 134 events remain in this subset, the result is therefore not surprising. A single event where none is expected, in a bin at a high magnitude, is not strong evidence for a deviation from a single power-law. We apply the Poisson GLM fit to four further subduction zone regions (Figure 3). All these catalogues also have histograms that lie within 95% confidence intervals of the GR distribution, with no evidence for characteristic events.

#### 5. Conclusions

[20] For all subduction zone earthquake catalogues considered in this paper, all histogram frequencies lie within Poisson counting fluctuations around a single power-law distribution. Other effects, for example, magnitude correlations or noise in the magnitude measurements, are too weak to be resolved. We find no evidence for rejecting the GR law in favour of the characteristic-type distribution.

[21] Inappropriate regression techniques provide biased estimates of histogram frequencies. This bias is amplified

by logarithmic transformation, such that the extrapolation of a power-law distribution is misleading, particularly for extreme events.

[22] Whilst taking larger samples improves the estimate of the distribution parameters, it cannot smooth the fluctuations in the tail of an unbounded power-law. Therefore, rejection of the GR distribution requires evidence of observations beyond counting errors.

[23] Some authors have noted that paleoseismological data record larger magnitude events than would be predicted by extrapolation of the GR model [e.g., *Speidel and Mattson*, 1997]. These outliers can be explained by a combination of the apparent positive bias for small samples at high magnitudes (even the best palaeoseismic records rarely contain more than 10 events), and the preferential preservation of such events in the geological record.

[24] In an era when earthquake science is required to advise civil society, our results have important implications for preparedness education in threatened centres of population. The characteristic model potentially constrains the location and magnitude of future events on a given fault segment, an appealing property that allows targeted preparedness strategies (e.g., retrofitting) to be put in place at lower cost. Here we have shown that it may be more prudent to consider a wider range of possibilities, including using Monte Carlo techniques to forecast hazard scenarios with a much wider range of possible consequences.

[25] **Acknowledgments.** MN was funded by EPSRC(GR/T11753/01), JG was funded by the EU TRIGS project (NEST-2005-PATH-COM-043386), and AB was funded by the EU NERIES project (RII3-CT-2006-026130) and collaboration NERC(NE/F01161X/1).

## References

- Aki, K. (1965), Maximum likelihood estimate of  $b$  in the formula  $\log N = a - bm$  and its confidence, *Bull. Earthquake Res. Inst. Univ. Tokyo*, **43**, 237–239.
- Borrero, J. C., K. Sieh, M. Chlieh, and C. E. Synolakis (2006), Tsunami inundation modeling for western Sumatra, *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 19,673–19,677, doi:10.1073/pnas.0604069103.
- Briggs, R. W., et al. (2006), Deformation and slip along the Sunda Megathrust in the Great 2005 Nias-Simeulue Earthquake, *Science*, **311**, 1897–1901, doi:10.1126/science.1122602.
- Charnes, A., E. L. Frome, and P. L. Yu (1976), The equivalence of generalised least squares and maximum likelihood estimates in the exponential family, *J. Am. Stat. Assoc.*, **71**, 169–171, doi:10.2307/2285762.
- Gutenberg, B., and C. Richter (1954), *Seismicity of the Earth and Associated Phenomena*, 2nd ed., 310 pp., Princeton Univ. Press, Princeton, N. J.
- Konca, A. O., et al. (2008), Partial rupture of a locked patch of the Sumatra megathrust during the 2007 earthquake sequence, *Nature*, **456**, 631–635, doi:10.1038/nature07572.
- Leonard, M. (2008), One hundred years of earthquake recording in Australia, *Bull. Seismol. Soc. Am.*, **98**, 1458–1470, doi:10.1785/0120050193.
- Liu, Y., A. Santos, S. M. Wang, Y. Shi, H. Liu, and D. A. Yuen (2007), Tsunami hazards along Chinese coast from potential earthquakes in South China Sea, *Phys. Earth Planet. Inter.*, **163**, 233–244, doi:10.1016/j.pepi.2007.02.012.
- Main, I. G., and P. W. Burton (1984), Information-theory and the earthquake frequency-magnitude distribution, *Bull. Seismol. Soc. Am.*, **74**, 1409–1426.
- Main, I. G., L. Li, J. McCloskey, and M. Naylor (2008), Effect of the Sumatran mega-earthquake on the global magnitude cut-off and event rate, *Nature Geosci.*, **1**, 142, doi:10.1038/ngeo141.
- McCloskey, J., A. Antonioli, A. Piatanesi, K. Sieh, S. Steacy, S. Nalbant, M. Cocco, C. Giunchi, J. Huang, and P. Dunlop (2008), Tsunami threat in the Indian Ocean from a future megathrust earthquake west of Sumatra, *Earth Planet. Sci. Lett.*, **265**, 61–81, doi:10.1016/j.epsl.2007.09.034.
- Polat, O., E. Gök, and D. Yilmaz (2008), Earthquake hazard of the Aegean Extension Region (west Turkey), *Turk. J. Earth Sci.*, **17**, 593–614.
- Sandri, L., and W. Marzocchi (2007), A technical note on the bias in the estimation of the  $b$ -value and its uncertainty through the least squares technique, *Ann. Geophys.*, **50**, 329–339.
- Sayil, N., and I. Osmansahim (2008), An investigation of seismicity for western Anatolia, *Nat. Hazards*, **44**, 51–64, doi:10.1007/s11069-007-9141-2.
- Schwartz, D. P., and K. J. Coppersmith (1984), Fault behavior and characteristic earthquakes: Examples from the Wasatch and San-Andreas fault zones, *J. Geophys. Res.*, **89**, 5681–5698, doi:10.1029/JB089iB07p05681.
- Sieh, K., D. H. Natawidjaja, A. J. Meltzner, C. C. Shen, H. Cheng, K. S. Li, B. W. Suwargadi, J. Galetzka, B. Philiposian, and R. L. Edwards (2008), earthquake supercycles inferred from sea-level changes recorded in the corals of west Sumatra, *Science*, **322**, 1674–1678, doi:10.1126/science.1163589.
- Speidel, D. H., and P. H. Mattson (1997), Problems for probabilistic seismic hazard analysis, *Nat. Hazards*, **16**, 165–179, doi:10.1023/A:1007976907691.
- Wesnowsky, S. G. (1994), The Gutenberg-Richter or characteristic earthquake distribution, which is it?, *Bull. Seismol. Soc. Am.*, **84**, 1940–1959.
- A. F. Bell, J. Greenough, I. G. Main, and M. Naylor, School of GeoSciences, University of Edinburgh, Edinburgh EH9 3JW, UK. (mark.naylor@ed.ac.uk)
- J. McCloskey, School of Environmental Sciences, University of Ulster, Ulster BT52 1SA, UK.